Hedging with Forecasting: A State–Space Approach to Modeling Vector-Valued Time Series

Tomislav Vukina

INTRODUCTION

The behavior of commodity markets has been the subject of prolonged research and controversy. Much of this controversy focuses on the behavior of futures prices over time. The fact that the application of certain technical trading systems can lead to positive returns leads some researchers to conclude that the behavior of futures prices is nonrandom. However, a number of economists believe that there are no systematic patterns to futures price behavior, and that expected profits cannot be increased by application of "... chart or any other esoteric devices of magic or mathematics" (Samuelson, 1965).

Studies of the behavior of futures prices are primarily concerned with pricing efficiency and hedging opportunities. A price efficient market is one that fully reflects all available information (Fama, 1970). An efficient market can be defined also as one in which a speculator can not earn an above-normal return (Panton, 1980). Much of the work in pricing efficiency of commodity futures markets can be classified in one of three areas: First, numerous studies test the random walk model [e.g., Leuthold (1972); Stevenson and Bear (1970)]. In general, the evidence is consistent with random behavior in futures prices. An efficient commodity market implies that futures prices over the life of a contract represent a martingale process. The random walk (martingale) hypothesis postulates that the expected futures price increments are not significantly different from zero, and futures price increments are independent, and thus, uncorrelated over time. Based on the several statistical tests of serial correlation applied to a total of 464 commodity contracts using daily prices, Cargill and Rausser (1975) reject the random walk model as a nonrealistic description of commodity markets. They conclude, however, that failure of the martingale property to hold is not an evidence of market inefficiency, and suggest that other more complicated martingale models may be more consistent with the observed data. The simple random walk is only a special case of the much broader Samuelson (1965) capital market model.

Tomislav Vukina is postdoctoral fellow at the Department of Resource Economics, University of Rhode Island, and research associate with the Institute for Development and International Relations, University of Zagreb, Croatia.

The Journal of Futures Markets, Vol. 12, No. 3, 307–327 (1992)

CCC 0270-7314/92/030307-21$04.00

Second, several studies apply mechanical filters to futures prices (usually based on the theory of normal backwardation) to determine whether profits can be realized [e.g., Gray (1961); Cootner (1960); Carter (1984)]. A major implication of the theory of normal backwardation is that a strategy of maintaining a long position in the futures market should earn positive profits over time. The results are ambiguous.

The third test of the efficient market hypothesis is the examination of the forecasting ability of futures markets. Futures prices are interpreted as reflecting the consensus of what traders expect the price to be at a certain time in the future, given the information currently available. Tomek and Gray (1970) view the springtime prices of the harvest-time futures as a forecast of the forthcoming harvest-time price. Kofi (1973) demonstrates that futures markets perform their forward pricing function very well, and that predictive reliability of a futures market improves as more accurate information on supply and demand becomes available. Both studies find that the forward pricing function of futures markets is more reliable for continuous than for discontinuous inventory models. Leuthold (1974) subsequently investigates the forecasting efficiency of futures markets using a similar approach of regressing cash prices on lagged futures prices. In his study, cash cattle prices are found to be more accurate predictors of subsequent cash cattle price than are the futures prices for distant contracts. The ability of futures prices to anticipate subsequent cash prices appears to be less efficient as time progresses. Martin and Garcia (1981) conclude that the livestock futures market has clearly not performed the forecasting function well. Just and Rausser (1981) address the question of whether futures markets are more or less accurate than the large-scale econometrically based forecasts. They find the evidence to be not overwhelmingly in favor of either.

The purpose of producer hedging is defined traditionally as a means to minimize possible revenue losses associated with adverse cash price changes. Alternative motives for hedging arise from either anticipating favorable basis change or attempts to diversify a portfolio of assets. One of the theoretical approaches to hedging most frequently used is the portfolio analysis [e.g., McKinnon (1967); Ward and Fletcher (1971); Rutledge (1972); Heifner (1972)]. These studies rely on traditional measures of expected returns and risk, i.e., the mean and variance of a series of past years prices to determine an optimal portfolio. As noted by Peck (1975), traditional measures of return and risk impose a long-run view, whereas futures markets operate in an essentially short-run horizon. With the use of price forecasts, Peck developed a hedging strategy which reduced risk by lowering producers exposure to unpredictable price variation. Brandt (1985) showed how both producers and first-handlers can reduce the risk of unfavorable price fluctuations by combining the information from forecasting models with a selective hedging strategy.

The central hypothesis of this article is that potential positive returns are available to traders who are able to outperform other traders in predicting price movements. The objective of the study is to show how price forecasts can be effectively used to formulate a hedging strategy which will increase profits from hedging. Cash and futures prices, or, alternatively, basis, are forecasted using an approach to modeling vector-valued time series based on linear systems theory. Hedging strategy based on price signals from these forecasts is described, followed by an overview of state-space procedure for modeling the vector-valued stationary time series. Finally, an example of the selective hedging strategy based on price forecasts applied to the soybean meal inventory is presented, and the results are compared to those obtained by routine hedging and cash marketing strategies.

METHODOLOGICAL FRAMEWORK

While examining the forecasting ability of futures markets, most previous studies compare cash prices with lagged futures prices. In this study, the forecasting ability of futures prices is exploited differently. As new information emerges or events occur, futures and cash prices are affected simultaneously. However, this does not imply that the previous prices were in error. It means only that they were based on a different information set than that subsequently available. Therefore, this approach uses contemporaneous cash and futures prices, estimates the parameters of the multivariate time series state-space model, and then generates out-of-sample forecasts. Thus, rather than using futures market prices to forecast future cash prices, the proposed method exploits dynamic relationship between contemporaneous time series of cash and futures prices, and forecasts them simultaneously.

It has been shown that the routine hedging strategy (forward pricing of the entire cash position) reduces risk [Peck (1975)]. However, this strategy frequently results in foregone potential positive returns from the correctly predicted movements in prices. A more sophisticated strategy, which incorporates the information from forecasts, can conceptually both increase expected utility and reduce risk relative to strictly cash marketing or routine hedging.

Based on Brandt (1985), a simple rule for selective hedging is utilized. The producer (seller) hedges and stays hedged as long as the current futures market price is greater than the forecasted cash price, or lifts the hedge and remains unhedged as long as the cash price is forecasted to be higher than the current futures market price. However, any analysis which compares futures prices with cash prices must also adjust for the basis, particularly if the hedger has no intention of delivery. Basis, defined as the difference between the futures price and cash price, is traditionally more stable over time than cash or futures prices and is, therefore, easier to forecast. Additionally, hedging is not costless. The costs of hedging include broker's commissions and opportunity costs of capital tied up in the form of initial margins and margin calls. Therefore, as an operational rule for selective hedging, the producer (seller) hedges and stays hedged whenever:

$$F_t - (B^*_{t+1|t} + C_h) > P^*_{t+1|t} \tag{1}$$

Using the definition of the basis, expression (1) becomes:

$$F_t - [(F^*_{t+1|t} - P^*_{t+1|t}) + C_h)] - P^*_{t+1|t} > 0$$

$$F_t > F^*_{t+1|t} + C_h \tag{2}$$

where C_h is the cost of hedging, F_t is the period t quotation of the selected futures contract price, $F^*_{t+1|t}$ is the next period forecast of the same futures contract price, $P^*_{t+1|t}$ is the next period forecast of cash price, and $B^*_{t+1|t}$ is the next period basis forecast; all forecasts based on information available at time t. The intent is to evaluate, ex post, whether sellers (producers), by exercising the above selective hedging strategy, could receive a higher price for their commodity compared to sellers who use a routine hedging procedure or cash marketing strategy only.

Although analytically equal, expressions (1) and (2) can lead to three distinct forecasting models. To use hedging rule (1), forecasts of basis and cash price in period $(t + 1)$ are needed. Two-channel time series may consist of cash and futures prices to give the basis forecasts indirectly as the difference between futures price forecast and cash price forecast. Alternatively, historical data for a particular location basis can be calculated by subtracting cash from futures prices, and then modeled together with cash prices. Using this approach, basis forecasts are generated directly along with the cash price forecasts. Finally, the third alternative, based on expression (2), involves univariate time series model of futures market prices only.

FORECASTING WITH THE STATE-SPACE MODELS

An approach to modeling multivariate time series based on linear systems theory is provided by Aoki (1987, 1990). Several other studies use applications of the method, as well as theoretical extensions of the original state-space technique. Aoki and Havenner (1989b) provide a comprehensive review of the system-theoretic time series. In Aoki and Havenner (1989a), the relationship to two other methods for Markovian model construction is established. Havenner and Aoki (1988) give an instrumental variables interpretation of the linear system theory estimation. Havenner and Criddle (1989) present an application of state-space modeling of time series to California cattle inventories and beef prices. Cerchi and Havenner (1988) use the method for modeling stock prices. Dorfman and Havenner (1991) apply the state-space technique to the problem of modeling cyclical patterns in supply and demand for California olives.

Consider the problem of fitting an approximate model to T observations on the m series of a zero-mean, weakly stationary vector-valued Gaussian stochastic process, $\{y_t\}$, with the covariance sequence: $\Gamma_1 = E[y_{t+1}y_t']$. It is assumed that y_t is a finite-dimensional process in the sense that it can be realized by passing the white noise through a finite dimensional linear system:

$$\begin{aligned} x_{t+1|t} &= Ax_{t|t-1} + Be_t \\ y_t &= Cx_{t|t-1} + e_t \end{aligned} \tag{3}$$

where the first equation is called state equation; the second equation is called observation equation; and $(n * n)$ matrix A, $(m * n)$ matrix C, and $(n * m)$ matrix B are the coefficient matrices to be estimated. The input, $\{e_t\}$, is a white Gaussian process with: $E[e_t e_s'] = \delta_{ts}\Psi$, where δ_{ts} is Kronecker's delta defined by $\delta_{ts} = 1$ if $t = s$, and $\delta_{ts} = 0$ otherwise. x_t is an $(n * 1)$ vector of unobservable states that are minimal sufficient statistics (optimal carriers of information) for the history of the process y_t. The subscripts on x refer to the conditional expectation of x in the period of the first subscript, given the information set at the time of the second subscript.

The stochastic model in eq. (3) is called the innovations model. By definition, the state vector $x_{t|t-1}$ contains all relevant information for the past history of the series y_t up to time $t - 1$. To ensure that the next period state vector $x_{t+1|t}$ is a set of sufficient statistics for y_{t+1}, the new information contained in e_t (innovation vector) must be incorporated into the states. Matrix B accomplishes this task in a manner that preserves the minimal sufficiency of the states. The procedure for estimating the parameters of the state-space innovation model is explained in the Appendix.

With estimates of the system matrices A, B, and C of the innovation model, state and observation equations can be solved for in- and out-of-sample forecasts. One way to proceed is to construct one-step-ahead forecasts. The expected value of y_t in observation eq. (3) is:

$$E(y_t) = \hat{y}_t = Cx_t \tag{4}$$

since, based on the assumption that e_t is independent of x_t, the expected value of e_t is zero. Predicting e_t is equivalent to predicting a random selection from a distribution with zero-mean and finite variance. Upon substitution of the observation equation into the state equation, the state equation becomes:

$$x_{t+1} = (A - BC)x_t + By_t \tag{5}$$

Multiplying both sides by C, advancing the time in the previous equation by one, and combining the two expressions, gives the one-step-ahead forecast:

$$\hat{y}_{t+1} = C(A - BC)x_t + CBy_t \tag{6}$$

which can now be compared with the actual value of y_{t+1} to calculate the one-step-ahead forecast error, $\hat{e}_{t+1} = y_{t+1} - \hat{y}_{t+1}$. Next, this data point is added to the existing pool of data points to update the system matrices, and so on. Alternatively, the same set of estimates may be kept while additional data points are incorporated, or the system matrices may be updated each time after several new data points are presented to the model. Rather than using a sequence of k one-step-ahead forecasts, one k-step-ahead forecast may be constructed as follows:

$$\hat{y}_{t+k+1} = CA^k[(A - BC)x_t + By_t] \qquad k = 0, 1, \ldots \tag{7}$$

Clearly, if $k = 0$, the above equation reduces to a one-step-ahead forecast as in eq. (6).

Both approaches require an initial value of x_t to begin the solution. Its effect declines as t increases, provided that y_t is weakly stationary and that A is asymptotically stable. To estimate x_1, some scheme of smoothing or backforecasting (backcasting) may be used. The one employed here is based on Abraham and Ledolter (1983). State and observation equations from (3) have to be slightly modified to give:

$$\begin{aligned} \hat{y}_t &= \hat{C}\hat{x}_t + \mu \\ \hat{e}_t &= y_t - \hat{y}_t \qquad t = 1, 2, \ldots, T \\ \hat{x}_{t+1} &= \hat{A}\hat{x}_t + \hat{B}\hat{e}_t \end{aligned} \tag{8}$$

The system of equations in (8) can be solved forward in time, using the assumption that $x_1 = 0$, to obtain an estimate of x_{T+1}. Inverting the timing, the last equation in (8) becomes:

$$\hat{x}_t = \hat{A}\hat{x}_{t+1} + \hat{B}\hat{e}_{t+1} \qquad t = T, T-1, \ldots, 1 \tag{9}$$

which is then, together with the first two equations in (8), solved backward in time to obtain an estimate of x_1. With an estimate of x_1 the system in (8) is solved forward in time again to obtain forecasts $\hat{y}_1, \hat{y}_2, \ldots \hat{y}_T$. The last x_t obtained is again x_{T+1} (generally different from the previous one where the solution is started with $x_1 = 0$), which can be used to generate out-of-sample forecasts for which the actual values of y_t $(y_{T+1}, y_{T+2}, \ldots)$ are not yet observed. Clearly, in this situation, forecast error $\hat{e}_t$ can not be computed, and k-step-ahead forecasts can be generated using the formula obtained by substituting expression (5) into (7):

$$\hat{y}_{t+k+1} = CA^k x_{t+1} + \mu \tag{10}$$

and adding back the mean of the series μ.

PERFORMANCE OF THE FORECASTING MODELS

The soybean meal data for the analysis presented here are obtained from various issues of *The Wall Street Journal.* Data used cover the period between June 1, 1989 and May 31, 1991, for a total of 506 data points. The cash price refers to 44% protein grade meal and represents the midpoint of the observed daily closing market quotation in Decatur, Illinois, in dollars and cents per ton. The futures data are daily closing prices for the Chicago Board of Trade soybean meal December contract. The trading unit is 100 short tons, and prices are quoted in dollars and cents per ton, with a minimum fluctuation of \$0.10 per ton (\$10.00 per contract), basis f.o.b. railcars, bulk, Decatur, Illinois. There is only one deliverable grade of meal with a minimum of 44% protein content. The December contract is chosen because it is the most actively traded soybean meal contract for which the continuous series of prices exists throughout the year. The first trading day in December is used as the switching point between the expiring and the incoming December contract.

As mentioned earlier, the proposed methodological framework leads to three different forecasting models, two of them being multivariate (two-channel), and one being a univariate (scalar) time series model. In the first model cash price and futures price are modeled together ($m = 2$). In the second, cash and basis are modeled together, and the third models futures prices alone ($m = 1$).

The method briefly outlined in the Appendix applies only to stationary time series. Since economic data frequently have trend components due to inflation, growth, technological progress, etc., an important step in accurate modeling of time series is the transformation to stationarity. Available methods exist for stochastic and deterministic trend removals [Aoki and Havenner (1989b); Cerchi and Havenner (1988)]. However, in studies like the present one which is concerned only with short-run forecasting, the appropriate procedure should utilize shorter periods (two to three years) for estimation purposes to avoid possible structural shifts that might occur. Under such circumstances, the likelihood of having a time series dominated by a trend is considerably smaller. The visual inspection of the plots of cash price, futures price, and basis series does not indicate the presence of a trend. Moreover, the poles of the system, i.e., the eigenvalues of the transition matrix A, are in all models; and, all runs are estimated to be inside the unit circle (moduli less than 1), which confirms the validity of the assumption. Only eigenvalues of the matrix A of unit magnitude, or very close to it, require the implementation of procedures for trend removal.

All three models are initially estimated with the first 128 observations, and one-step-ahead-out-of-sample forecasts are generated using expression (10).[1] In the following step, an additional actual observation (number 129) is added to the existing pool of data points to update the system matrices, thus generating the new one-step-ahead forecast. The same procedure is repeated 378 times. Both the bivariate and the scalar model in each of the 378 rounds are estimated with the most distant prespecified autocovariance lag of $k = 2$. The decaying pattern of the diagonal matrix of singular values Σ suggest the number of states (the rank of the Hankel matrix) of $n = 2$. The resulting forecasts are plotted against the actual data in Figures 1–3.

To discuss alternative models it is necessary to have a criterion by which the performance of the predictors is measured. The usual criterion is the mean square error (MSE) or root mean square error (RMSE). However, to compare the models which attempt to model different processes (time series), the mean absolute percentage error (MAPE) is used, which is the measure of the average percentage deviation of the forecast from its corresponding actual value, and is therefore dimensionless. The model which generates forecasts with the smallest statistics is, according to these criteria, the most favorable one. The comparative forecasting statistics are presented in Table I. According to MAPE criterion, the best result is obtained with two-channel

[1]The program code written in MATLAB is available from the author upon request.

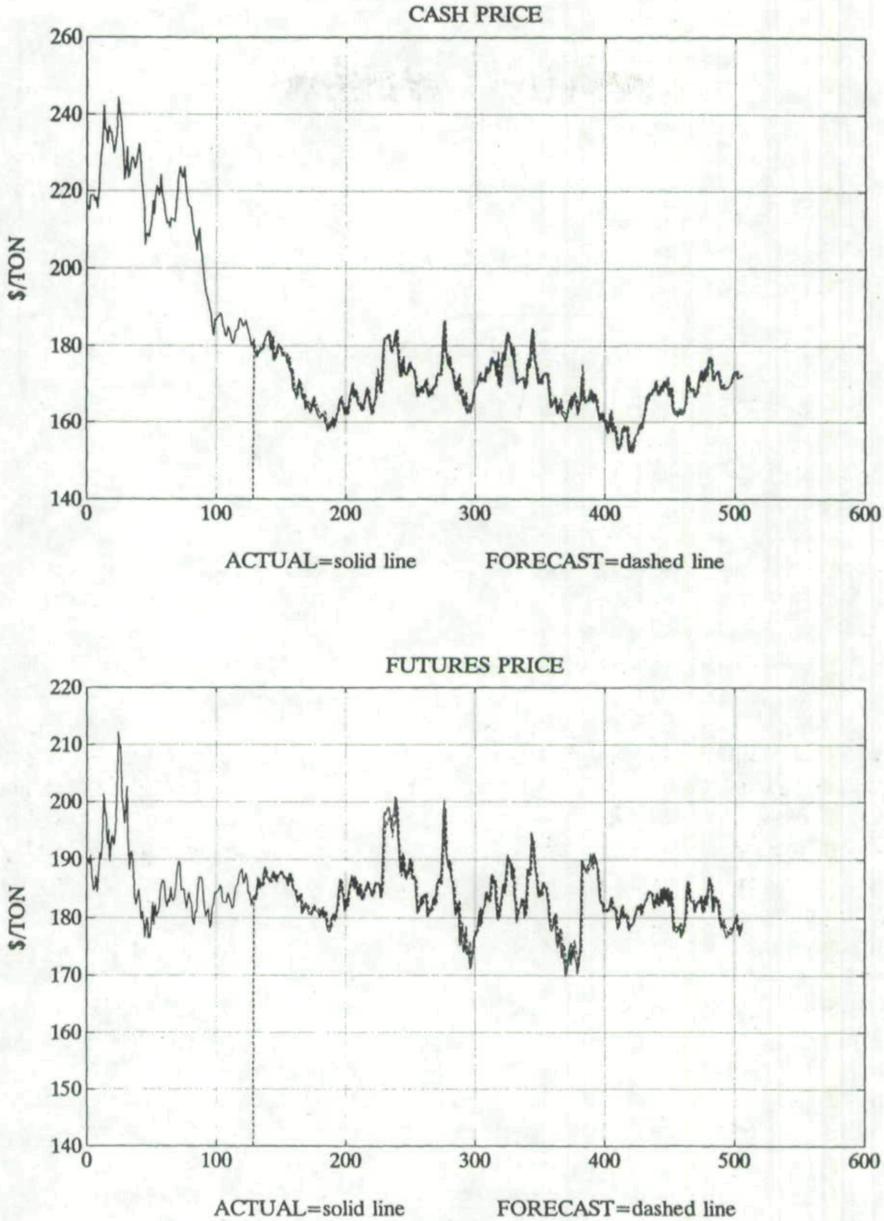

Figure 1

Two-channel (cash and futures price) model: daily soybean meal Decatur cash price and CBOT soybean meal futures price for the June 1, 1989 to May 31, 1991 period.

cash and futures model in the case of futures price series. The average percentage deviation of the forecast from its corresponding actual value is less than 0.8%. The forecasts of cash price are equally good in both multivariate models, with MAPE being less than 1%. The basis forecast yields more modest results, with a mean absolute percentage error close to 9%.

However, from the perspective of this study, a more valuable criterion for forecast evaluation is their usefulness for different hedging strategies. This issue is addressed next.

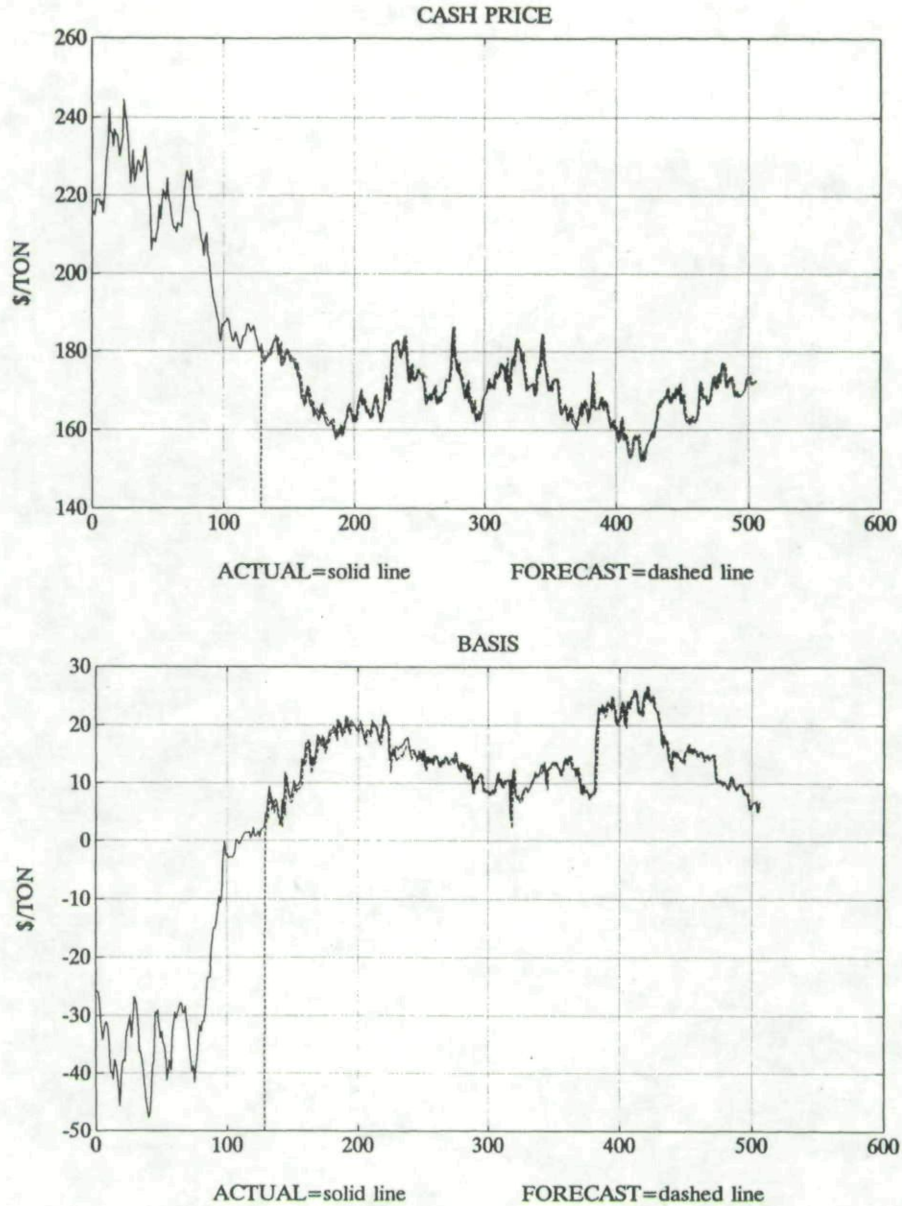

Figure 2
Two-channel (cash and basis) model: daily soybean meal Decatur cash price and Decatur basis for the June 1, 1989 to May 31, 1991 period.

COMBINING FORECASTS WITH HEDGING

Although the data on cash and futures soybean meal prices are available from June 1, 1989, the actual trading simulation starts on December 1, 1989 (the first 128 observations are used for initial estimation), and ends on May 31, 1991. The entire trading period is divided into six three-month long subperiods (see Tables II and III). At the end of each subperiod all outstanding cash and futures positions have to be liquidated. On the first trading day in each subperiod, a hypothetical feedstuffs dealer buys 100 tons of soybean meal at the prevailing spot market price, and sells its inven-

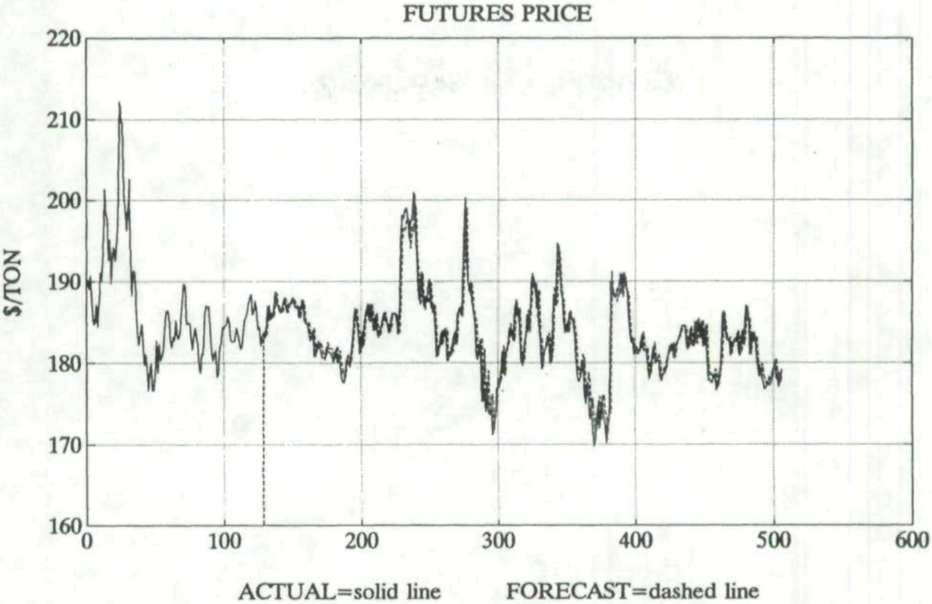

Figure 3
Scalar (futures price) model: daily CBOT soybean meal futures prices for the June 1, 1989 to May 31, 1991 period.

tory on the last trading day of the same period. To keep the problem tractable, the possibility of selling the inventory at any point in time between the two specified dates is ruled out. Since storage and handling costs are ignored, the difference in cash price between the two dates multiplied by the quantity (100 tons) is considered as profit from cash marketing.

During the period when a dealer is long in the cash market, the dealer faces three options: (i) to remain unhedged throughout the period; (ii) to hedge the entire inventory and remain hedged throughout the period (routine strategy); or (iii) hedge selectively using the operational rule specified in expression (1). For the sake of simplicity, it is further assumed that opportunity costs of capital (interest payments) are negligible and that hedging costs consist of only fixed costs per contract (brokerage fees). Because of the intraperiod price variation, margin calls present a distinct possibility and could potentially cause a cash flow problem for the hedger. However, since the likelihood of margin calls cannot a priori be considered any greater than the likelihood of margin surpluses, which would yield a favorable return to the hedger, the above assumption is probably not too restrictive.

If the dealer decides to remain unhedged (i), he assumes the entire risk of price variation, and his end-of-period total profit consists of profits from cash marketing only.

Routine hedging strategy (ii) involves selling short the entire inventory. The hedge is placed (one December contract is sold) on the first trading day of each subperiod, and lifted (one December contract is bought back) on the last trading day of the same period. Under the routine scenario, the profit per one round of hedging is equal to the difference between the futures market quotation on the day when the hedge is placed and the day when it is lifted, multiplied by the quantity hedged (1 contract = 100 tons), and corrected for the fixed cost of hedging per one contract, and per one roundturn.

Table I
COMPARATIVE FORECASTING STATISTICS

Model	Number of Observations: Initial Round	Number of Observations: Final Round	Number of One-Step-Ahead Forecasts	MSE	RMSE	MAPE
Two-channel ($m = 2, k = 2, n = 2$)	128	505	378			
Cash				4.9183	2.2177	0.9977%
Futures				4.5490	2.1328	0.7995%
Two-channel ($m = 2, k = 2, n = 2$)	128	505	378			
Cash				4.9229	2.2188	0.9978%
Basis				2.4291	1.5586	8.9552%
Scalar ($m = 1, k = 2, n = 2$)	128	505	378			
Futures				4.6465	2.1556	0.8069%

Table II
FORECAST-GENERATED HEDGING SIGNALS

	Cost = \$50.00 per Contract per Roundturn						
Period	No. of Decision Points	No. of HEDGE Signals	No. of Correct Signals	% of Correct Signals	No. of NO HEDGE Signals	No. of Correct Signals	% of Correct Signals
Two-channel model (cash + futures):							
Dec. 1, 1989–Feb. 28, 1990	60	0	0		60	31	51.67%
Mar. 1, 1990–May 31, 1990	63	24	13	54.17%	39	24	61.54%
Jun. 1, 1990–Aug. 31, 1990	64	9	7	77.78%	55	36	65.45%
Sep. 4, 1990–Nov. 30, 1990	62	6	5	83.33%	56	30	53.57%
Dec. 3, 1990–Feb. 28, 1991	60	14	11	78.57%	46	30	65.22%
Mar. 1, 1991–May 31, 1991	63	0	0		63	37	58.73%
Total	372	53	36	67.92%	319	188	58.93%
Two-channel model (cash + basis):							
Dec. 1, 1989–Feb. 28, 1990	60	0	0		60	31	51.67%
Mar. 1, 1990–May 31, 1990	63	24	13	54.17%	39	24	61.54%
Jun. 1, 1990–Aug. 31, 1990	64	9	7	77.78%	55	36	65.45%
Sep. 4, 1990–Nov. 30, 1990	62	6	5	83.33%	56	30	53.57%
Dec. 3, 1990–Feb. 28, 1991	60	14	11	78.57%	46	30	65.22%
Mar. 1, 1991–May 31, 1991	63	0	0		63	37	58.73%
Total	372	53	36	67.92%	319	188	58.93%
Scalar model (futures):							
Dec. 1, 1989–Feb. 28, 1990	60	0	0		60	31	51.67%
Mar. 1, 1990–May 31, 1990	63	19	10	52.63%	44	28	63.64%
Jun. 1, 1990–Aug. 31, 1990	64	11	9	81.82%	53	36	67.92%
Sep. 4, 1990–Nov. 30, 1990	62	12	8	66.67%	50	27	54.00%

Table II (continued)

Period	No. of Decision Points	No. of HEDGE Signals	No. of Correct Signals	% of Correct Signals	No. of NO HEDGE Signals	No. of Correct Signals	% of Correct Signals
	Cost = $50.00 per Contract per Roundturn						
Dec. 3, 1990–Feb. 28, 1991	60	14	11	78.57%	46	30	65.22%
Mar. 1, 1991–May 31, 1991	63	0	0		63	37	58.73%
Total	372	56	38	67.86%	316	189	59.81%

Period	No. of Decision Points	No. of HEDGE Signals	No. of Correct Signals	% of Correct Signals	No. of NO HEDGE Signals	No. of Correct Signals	% of Correct Signals
	Cost = $30.00 per Contract per Roundturn						
Two-channel model (cash + futures):							
Dec. 1, 1989–Feb. 28, 1990	60	23	15	65.22%	37	17	45.95%
Mar. 1, 1990–May 31, 1990	63	32	19	59.38%	31	20	64.52%
Jun. 1, 1990–Aug. 31, 1990	64	9	7	77.78%	55	34	61.82%
Sep. 4, 1990–Nov. 30, 1990	62	14	11	78.57%	48	27	56.25%

Dec. 3, 1990–Feb. 28, 1991	60	14	11	78.57%	46	30	65.22%
Mar. 1, 1991–May 31, 1991	63	0	0		63	33	52.38%
Total	372	92	63	68.48%	280	161	57.50%
Two-channel model (cash + basis):							
Dec. 1, 1989–Feb. 28, 1990	60	23	15	65.22%	37	17	45.95%
Mar. 1, 1990–May 31, 1990	63	32	19	59.38%	31	20	64.52%
Jun. 1, 1990–Aug. 31, 1990	64	9	7	77.78%	55	34	61.82%
Sep. 4, 1990–Nov. 30, 1990	62	14	11	78.57%	48	27	56.25%
Dec. 3, 1990–Feb. 28, 1991	60	14	11	78.57%	46	30	65.22%
Mar. 1, 1991–May 31, 1991	63	0	0		63	33	52.38%
Total	372	92	63	68.48%	280	161	57.50%
Scalar model (futures):							
Dec. 1, 1989–Feb. 28, 1990	60	0	0		60	28	46.67%
Mar. 1, 1990–May 31, 1990	63	20	12	60.00%	43	25	58.14%
Jun. 1, 1990–Aug. 31, 1990	64	13	10	76.92%	51	33	64.71%
Sep. 4, 1990–Nov. 30, 1990	62	15	10	66.67%	47	25	53.19%
Dec. 3, 1990–Feb. 28, 1991	60	14	11	78.57%	46	30	65.22%
Mar. 1, 1991–May 31, 1991	63	5	4	80.00%	58	33	56.90%
Total	372	67	47	70.15%	305	174	57.05%

Table III

PROFITS FROM HEDGING ONE CONTRACT (100 TONS) OF CBOT SOYBEAN MEAL FOR THE PERIOD DECEMBER 1, 1989 TO MAY 31, 1991

			Selective Hedging					
			Two-Channel Model (Cash + Futures)		Two-Channel Model (Cash + Basis)		Scalar Model (Futures)	
Period	Cash Marketing Profits	Routine Hedging Profits	Profits	No. of Turns	Profits	No. of Turns	Profits	No. of Turns
Cost = $50.00 per contract per roundturn:								
Dec. 1, 1989–Feb. 28, 1990	−$1850	$340	$0	0	$0	0	$0	0
Mar. 1, 1990–May 31, 1990	$1200	−$710	$730	1	$730	1	$870	2
Jun. 1, 1990–Aug. 31, 1990	−$75	$400	$1580	1	$1580	1	$2020	3
Sep. 4, 1990–Nov. 30, 1990	−$525	$880	$1060	1	$1060	1	$440	3
Dec. 3, 1990–Feb. 28, 1991	−$450	$740	$680	1	$680	1	$680	1
Mar. 1, 1991–May 31, 1991	$300	$530	$0	0	$0	0	$0	0
Total	−$1400	$2180	$4050	4	$4050	4	$4010	9
Percent of routine hedge		100.00%	185.78%		185.78%		183.94%	
Cost = $30.00 per contract per roundturn:								
Dec. 1, 1989–Feb. 28, 1990	−$1850	$360	$360	1	$360	1	$0	0
Mar. 1, 1990–May 31, 1990	$1200	−$690	$1140	2	$1140	2	$1260	2
Jun. 1, 1990–Aug. 31, 1990	−$75	$420	$1600	1	$1600	1	$1930	3
Sep. 4, 1990–Nov. 30, 1990	−$525	$900	$1180	2	$1180	2	$320	4
Dec. 3, 1990–Feb. 28, 1991	−$450	$760	$700	1	$700	1	$700	1
Mar. 1, 1991–May 31, 1991	$300	$550	$0	0	$0	0	$630	2
Total	−$1400	$2300	$4980	7	$4980	7	$4840	12
Percent of routine hedge		100.00%	216.52%		216.52%		210.43%	

Selective hedging strategy (iii) also involves selling short the entire inventory. However, the hedge is placed only when the action is signaled by the forecasts, i.e., when the condition specified in (1) is fulfilled. More specifically, every trading day the parameters of the model are estimated, the prices for the next day (cash, futures, basis) are forecasted, and the decision whether to hedge or not is made. The following day, the previous day prices are added to the pool of the existing data, the parameters of the model are reestimated and the new one-step-ahead forecasts are made together with the new decision.

The performance evaluation of the forecasting approach to hedging under two different cost scenarios for the period from December 1, 1989 to May 31, 1991, is summarized in Table II. Initially, trading is simulated with a fixed cost of \$50.00 per contract per roundturn, and then the entire exercise is repeated using a cost of \$30.00. Notice, however, that once the hedge is in place, it is costless to continuously remain hedged in the same number of contracts afterward. In this case, in the evaluation of either condition (1) or condition (2), $C_h = 0$ is required. Only after a disruption of the continuous sequence of HEDGE signals occurs (hedge is lifted), the establishment of the new position in the futures market requires additional expenditure. Or, if the position in the futures market is already established, the cost of retaining the same position continuously through time is zero. Costs of hedging in this simplified model are associated only with changing the position in the futures market or entering it for the first time.

Measured by the percentage of correct signals, the results indicate no significant difference between forecasting models. In fact, both multivariate models yield exactly identical results, and, measured by the combined percentage of correct HEDGE and NO HEDGE signals, slightly outperform the scalar model. The hedging cost reduction from \$50.00 to \$30.00 per contract per roundturn, resulted in an increased number of HEDGE signals. As seen from expressions (1) or (2), the reduction in hedging cost affects the signal-generating formulae by softening the restrictiveness of the conditions for positive signals.

Measured by the amount of trading profits generated, both multivariate models perform equally well, and generate higher profits compared to the scalar model. As seen from Table III, the profits generated by either one of the multivariate models are 86% larger than those obtained by the routine strategy. The scalar model generates profits 84% in excess of those obtained by routine hedging. The number of roundturns per period is relatively small. In fact, for both multivariate models, the futures market is entered only once in four subperiods and zero times in the remaining two. In the case of one roundturn per period, the number of days a hypothetical trader is hedged corresponds to the number of HEDGE signals from Table II. For example, it turns out that in the \$50.00 cost scenario, during the period from March 1, 1990 to May 31, 1990, the trader using forecasts from the two-channel cash and futures model enters the futures market once and remains hedged for 24 consecutive trading days. In the case of the multiple entry-exit, as in the scalar model, the average duration of the short hedge can be calculated as the ratio between the number of HEDGE signals and the number of roundturns.

The hedging cost reduction causes the difference between the selective and the routine strategy to be even more pronounced. In the \$30.00 cost scenario, the results show that the hypothetical feedstuffs dealer can increase his profits by 117% above the routine strategy by hedging selectively using either one of the vector-valued time series forecasting models. The application of the scalar forecasting model earns him profits 110% over the routine hedging. As intuitively expected, the reduction in cost contributed also to the increased number of roundturns per period.

CONCLUSIONS

Numerous simplifying assumptions are made in this study. A simple decision rule is used; no attempts are made to determine the optimal hedging ratio in the static framework, let alone in the intertemporal framework; and opportunity costs of capital and cash flow problems are ignored. Also, no attempt is made to incorporate hedging strategy into the broader framework of optimal inventory management. Additional work of this type is needed to find out whether these modifications would substantially alter the basic conclusions.

The main goal of the study is to show the usefulness of one of the state-of-the-art time series forecasting techniques in marketing agricultural commodities for which established futures markets exist. The study illustrates how agricultural commodity dealers (producers, sellers), even in the presence of substantial hedging costs, can significantly increase their profits by combining the information from the forecasting models with an appropriate hedging strategy.

The study also shows the superiority of the multivariate time series approach over scalar time series modeling. The forecasts generated by the models which exploit a dynamic relationship between contemporaneous time series of cash and futures prices (or cash and basis) and forecast them simultaneously, when combined with a simple decision rule, result in larger profits than those generated by the competing scalar forecasting model.

Since the basis is traditionally more stable over time than cash or futures prices, most researchers feel that it should be easier to forecast the basis than either cash or futures prices. However, the results of this study do not confirm this hypothesis. First, the two-channel model, with cash and basis modeled together, do not generate more profits than the cash and futures prices model. Second, the mean absolute percentage error for the one-step-ahead out-of-sample forecasts of the basis of 9% is significantly larger than MAPEs for other forecasted series which are all under 1%.

Finally, the forecasts from the state-space models of time series are relatively easy to generate once the computer program is written. Decision makers could generate their own forecasts using this type of model on a microcomputer, without the necessity to understand relatively complicated sections from the underlying linear systems theory.

Appendix

To get zero-mean time series and autocovariances of the same order of magnitude, the original time series has to be centered and scaled. This is accomplished by subtracting the mean of the time series from each individual observation and dividing it by the series standard deviation. In terms of the original data, the state-space innovation model (3) becomes:

$$\begin{aligned} x_{t+1|t} &= Ax_{t|t-1} + BD^{-1}De_t \\ D(y_t - \mu) &= DCx_{t|t-1} + De_t \end{aligned} \tag{A-1}$$

where μ is the m series mean vector and D a diagonal $(m * m)$ matrix of standard deviations. It is convenient to stack the observations, temporally defining the following vectors:

$$y_t^+ = \begin{bmatrix} y_t \\ y_{t+1} \\ \cdot \\ \cdot \\ \cdot \\ y_{t+k-1} \end{bmatrix} \qquad y_t^- = \begin{bmatrix} y_t \\ y_{t-1} \\ \cdot \\ \cdot \\ \cdot \\ y_{t-k+1} \end{bmatrix} \qquad y_{t-1}^- = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \cdot \\ \cdot \\ \cdot \\ y_{t-k} \end{bmatrix} \tag{A-2}$$

Now consider the $(mk * mk)$ matrix H formed by arranging the first k autocovariance matrices of observations into a block-band counterdiagonal matrix called the Hankel form:

$$H = E(y_t^+ y_{t-1}^{-'}) = E \begin{bmatrix} (y_t y_{t-1}') & (y_t y_{t-2}') & \cdots & (y_t y_{t-k}') \\ (y_{t+1} y_{t-1}') & (y_{t+1} y_{t-2}') & \cdots & (y_{t+1} y_{t-k}') \\ \cdot & & & \\ \cdot & & & \\ (y_{t+k-1} y_{t-1}') & (y_{t+k-1} y_{t-2}') & \cdots & (y_{t+k-1} y_{t-k}') \end{bmatrix} \tag{A-3}$$

The sample estimate of the lag l aultocovariance matrix is:

$$\hat{\Gamma}_l = \frac{1}{T} \sum_{t=1}^{T-1} y_{t+1} y_t' \qquad l = 0, 1, 2, \ldots, 2k \tag{A-4}$$

and the Hankel matrix becomes:

$$\hat{H} = \begin{bmatrix} \hat{\Gamma}_1 & \hat{\Gamma}_2 & \cdots & \hat{\Gamma}_k \\ \hat{\Gamma}_2 & \hat{\Gamma}_3 & \cdots & \hat{\Gamma}_{k+1} \\ \cdot & & & \\ \cdot & & & \\ \hat{\Gamma}_k & \hat{\Gamma}_{k+1} & \cdots & \hat{\Gamma}_{2k-1} \end{bmatrix} \tag{A-5}$$

where $(2k - 1)$ is the maximally distant autocovariance lag to be modeled.

The Hankel matrix (A-3) can be regarded as the $(mk * mk)$ upper left corner of an infinite by infinite Hankel matrix over all time. Assuming that k bounds the model order (i.e., for sufficiently large k), the theoretical rank of the Hankel matrix in (A-3) is the same as the rank of its underlying doubly infinite population counterpart; and further, by the Kronecker theorem, this rank represents the number of states required to characterize the past history of the system [Kailath (1980)].

A computationally accurate method of determining the rank of Hankel matrix is provided by the singular value decomposition (SVD):

$$H = U \, \Sigma \, V'$$

where Σ is diagonal matrix of singular values arranged in descending order:

$$\sum = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_{mk}); \ \sigma_1 \geq \sigma_2 \geq \ \ldots \geq \sigma_{mk}$$

and the associated vectors are orthonormal:

$$U'U = V'V = I$$

The essence of the procedure is to approximate the space spanned by the Hankel matrix of autocovariances by the subspace spanned by those singular vectors associated with nonzero singular values. This can be viewed as an approximation of the observed autocovariance sequence arranged in $\hat{H}$ by the rank $\tilde{n}$ matrix $\hat{H}_{\tilde{n}}$ with the following singular value decomposition:

$$\hat{H}_{\tilde{n}} = \hat{U}_{\tilde{n}} \hat{\Sigma}_{\tilde{n}} \hat{V}'_{\tilde{n}} \tag{A-6}$$

where the subscripts indicate that only the largest $\tilde{n}$ singular values are included in the approximation. Estimating model parameters with truncating the Hankel matrix by excluding those singular values and singular vectors beyond $\tilde{n}$ is analogous to the usual exclusion restrictions of other time series modeling procedures. The difference is that the exclusions are imposed on the singular values rather than on lagged y's and errors.

The covariance of the states with the observations is defined as follows:

$$\Omega = E(X_{t+1} y'_t)$$

Postmultiplying the observation equation in (3) by successive lags of y and taking expectations establishes that:

$$\Gamma_l = C A^{l-1} \Omega; \quad l = 1, 2, \ldots, 2k - 1$$

and the Hankel matrix then becomes:

$$H = \begin{bmatrix} C\Omega & CA\Omega & \cdots & CA^{k-1}\Omega \\ CA\Omega & CA^2\Omega & \cdots & CA^k\Omega \\ \vdots & & & \\ CA^{k-1}\Omega & CA^k\Omega & \cdots & CA^{2k-2}\Omega \end{bmatrix} \tag{A-7}$$

Recognition of this pattern allows another factorization of the Hankel matrix into its observability and controllability matrices, O and K:

$$H = OK = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{k-1} \end{bmatrix} [\Omega \; A\Omega \; A^2\Omega \cdots A^{k-1}\Omega] \tag{A-8}$$

Note that the $(m * n)$ coefficient matrix C is visible in the observability matrix, i.e., $J'O = C$; where $J' = [I\,0 \ldots 0]$; and I is an $(m * m)$ identity matrix. Similarly, the $(n * m)$ covariance matrix, Ω, is visible in controllability matrix, i.e., $KJ = \Omega$. Therefore, the first block row of the Hankel matrix can be defined as:

$$\begin{aligned} H_{1\cdot} &= J'OK = CK \\ &= [\Gamma_1 \; \Gamma_2 \cdots \Gamma_k] \end{aligned} \tag{A-9}$$

Similarly, the first block column (m columns) of H can be defined as:

$$H_{\cdot 1} = OKJ = O\Omega = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \vdots \\ \Gamma_k \end{bmatrix} \tag{A-10}$$

Next, define a time-shifted autocovariance matrix:

$$\overleftarrow{H} = OAK = \begin{bmatrix} C \\ CA \\ \cdot \\ \cdot \\ CA^{k-1} \end{bmatrix} [A][\Omega\ A\Omega \cdots A^{k-1}\Omega]$$

$$= \begin{bmatrix} \Gamma_2 & \Gamma_3 & \cdots & \Gamma_{k+1} \\ \Gamma_3 & \Gamma_4 & \cdots & \Gamma_{k+2} \\ \cdot & & & \\ \cdot & & & \\ \Gamma_{k+1} & \Gamma_{k+2} & \cdots & \Gamma_{2k} \end{bmatrix} \quad \text{(A-11)}$$

Two different factorizations of the Hankel matrix [eqs. (A.6) and (A.8)] can be equated, recalling that the $\tilde{n}$ subscript indicates approximation based on the included singular values, which gives:

$$\hat{H}_{\tilde{n}} = \hat{U}_{\tilde{n}} \hat{\sum}_{\tilde{n}}^{\frac{1}{2}} \hat{\sum}_{\tilde{n}}^{\frac{1}{2}} \hat{V}_{\tilde{n}}' = \hat{O}_{\tilde{n}} \hat{K}_{\tilde{n}} \quad \text{(A-12)}$$

$$\hat{O}_{\tilde{n}} = \hat{U}_{\tilde{n}} \hat{\sum}_{\tilde{n}}^{\frac{1}{2}}$$

$$\hat{K}_{\tilde{n}} = \hat{\sum}_{\tilde{n}}^{\frac{1}{2}} \hat{V}_{\tilde{n}}'$$

Let:

$$\hat{O}_{\tilde{n}}^{L} = (\hat{O}_{\tilde{n}}' \hat{O}_{\tilde{n}})^{-1} \hat{O}_{\tilde{n}}'$$

be the left generalized inverse of the observability matrix so that $\hat{O}_{\tilde{n}}^{L} \hat{O}_{\tilde{n}} = \mathrm{I}$. Since $\Sigma_{\tilde{n}}$ is diagonal matrix and it equals its transposition, and $\hat{U}_{\tilde{n}}' \hat{U}_{\tilde{n}} = I$, a straightforward substitution gives the left singular value generalized inverse of $\hat{O}_{\tilde{n}}$:

$$\hat{O}_{\tilde{n}}^{L} = \hat{\sum}_{\tilde{n}}^{-\frac{1}{2}} \hat{U}_{\tilde{n}}' \quad \text{(A-13)}$$

Analogously, the right singular value generalized inverse of the controllability matrix is:

$$\hat{K}_{\tilde{n}}^{R} = \hat{V}_{\tilde{n}} \hat{\sum}_{\tilde{n}}^{-\frac{1}{2}} \quad \text{(A-14)}$$

Using generalized inverses obtained in (A-13) and (A-14), the solution to (A-9), (A-10), and (A-11) provides estimates of the matrices C, Ω, and A:

$$\hat{C} = \hat{H}_{1\cdot} \hat{K}_{\tilde{n}}^{R} \quad \text{(A-15)}$$

$$\hat{\Omega} = \hat{O}_{\tilde{n}}^{L} \hat{H}_{\cdot 1} \quad \text{(A-16)}$$

$$\hat{A} = \hat{O}_{\tilde{n}}^{L} \overleftarrow{H} \hat{K}_{\tilde{n}}^{R} \quad \text{(A-17)}$$

To develop an estimate of the Kalman filter matrix B, one must jointly solve three matrix equations that are derived from the state covariance matrix $P = E\ (X_{t+1|t}\ X_{t+1|t}')$; the error covariance matarix $\Psi = E(e_t e_t')$; and the covariance of the states with the data, $\Omega = E\ (X_{t+1|t} y_t')$, along with the unconditional data covariance matrix $\Gamma_0 = E(y_t y_t')$. Using the property of orthogonality of the innovations and the current states: $E\ (e_t X_{t|t-1}') = 0$; and taking expectations of the state equation in (3), with itself transposed, gives:

$$P = A\,P\,A' + B\,\Psi B'$$

Similarly, taking expectations of the observation equation in (3), with itself transposed, gives:

$$\Gamma_0 = C\,P\,C' + \Psi$$

As evident from above, Ψ is positive definite whenever the model adds information, i.e., when states explain some part of the unconditional data covariance so that $(\Gamma_0 > C\ P\ C')$. This is not, however, imposed on the sample estimates, requiring that the estimated Ψ matrix be checked for the positive definiteness. Finally, taking expectations of the state equation with the observation equation transposed gives:

$$\Omega = A\ P\ C' + B\ \Psi$$

The above equations can be solved simultaneously for the estimates of P, Ψ, and B. By slight rearrangement and direct substitution, one obtains a special form of matrix Riccati equation:

$$P = APA' + (\Omega - APC')(\Gamma_0 - CPC')^{-1}(\Omega - APC')' \tag{A-18}$$

which can be solved for P using iterative as well as noniterative solution algorithms. In solving the Riccati equation, a nonrecursive solution algorithm derived by Vaccaro and Li (1987) is followed. Once the estimate of P is obtained, the other two estimates follow directly:

$$\hat{\Psi} = \hat{\Gamma}_0 - \hat{C}\hat{P}\hat{C}' \tag{A-19}$$

$$\hat{B} = (\hat{\Omega} - \hat{A}\hat{P}\hat{C}')\hat{\Psi}^{-1} \tag{A-20}$$

Bibliography

Abraham, B., and Ledolter, J. (1983): *Statistical Methods for Forecasting,* New York: Wiley.

Aoki, M. (1987): *State Space Modeling of Time Series,* Berlin: Springer Verlag.

Aoki, M. (1990): *State Space Modeling of Time Series* (2nd ed.). Berlin: Springer Verlag.

Aoki, M., and Havenner, A. (1989a): "A Method for Approximate Representation of Vector-Valued Time Series and Its Relation to Two Alternatives," *Journal of Econometrics,* 42:181–199.

Aoki, M., and Havenner, A. (1989b): "State Space Modeling of Multiple Time Series," Agricultural Economics Working Paper 89-9, University of California, Davis.

Brandt, J. A. (1985): "Forecasting and Hedging: An Illustration of Risk Reduction in the Hog Industry," *American Journal of Agricultural Economics,* 67:24–31.

Cargill, T., and Rausser, G. C. (1975): "Temporal Price Behavior in Commodity Futures Markets," *Journal of Finance,* 30:1043–1053.

Carter, C. A. (1984): "An Evaluation of Pricing Performance and Hedging Effectiveness of the Barley Futures Market," *Western Journal of Agricultural Economics,* 9:1–13.

Cerchi, M., and Havenner, A. (1988): "Cointegration and Stock Prices: The Random Walk on Wall Street Revisited," *Journal of Economic Dynamics and Control,* 12:333–346.

Cootner, P. H. (1960): "Returns to Speculators: Telser Versus Keynes," Journal of Political Economy, 68:396–404.

Dorfman, J. H., and Havenner, A. (1991): "State Space Modeling of Cyclical Supply, Seasonal Demand, and Agricultural Inventories" *American Journal of Agricultural Economics,* 73:829–840.

Fama, E. F. (1970): "Efficient Capital Markets: A Review of Theory and Empirical Work," *The Journal of Finance,* 25:383–423.

Gray, R.W. (1961): "The Search for a Risk Premium," *Journal of Political Economy,* 69:250–260.

Havenner, A., and Aoki, M. (1988): "An Instrumental Variables Interpretation of Linear Systems Theory Estimation," *Journal of Economic Dynamics and Control,* 12:49–54.

Havenner, A. M., and Criddle, K. R. (1989): "System Theoretic Time Series: An Application to Inventories and Prices of California Range Cattle," Computers and Mathematics with Applications, 17:1177–1187.

Heifner, R. J. (1972): "Optimal Hedging Levels and Hedging Effectiveness in Cattle Feeding," *Agricultural Economics Research,* 24:25–36.

Just, R. E., and Rausser, G. C. (1981): "Commodity Price Forecasting with Large-Scale Econometric Models and the Futures Market," *American Journal of Agricultural Economics,* 63:197–208.

Kailath, T. (1980): Linear System Theory, Englewood Cliffs, NJ: Prentice-Hall.

Kofi, T. A. (1973): "A Framework for Comparing the Efficiency of Futures Markets," *American Journal of Agricultural Economics,* 55:584–594.

Leuthold, R. M. (1972): "Random Walk and Price Trends: The Live Cattle Futures Market," *The Journal of Finance,* 27:879–889.

Leuthold, R. M. (1974): "The Price Performance on the Futures Market of a Nonstorable Commodity: Live Beef Cattle," *American Journal of Agricultural Economics,* 56:271–279.

Martin, L., and Garcia, P. (1981): "The Price-Forecasting Performance of Futures Markets for Live Cattle and Hogs: A Disaggregated Analysis," *American Journal of Agricultural Economics,* 63:209–215.

McKinnon, R. I. (1967): "Futures Markets, Buffer Stocks and Income Stability for Primary Producers," *Journal of Political Economy,* 75:844–861.

Panton, D. B. (1980): "A Semi-Strong Form Evaluation of the Efficiency of the Hog Futures Market: Comment," *American Journal of Agricultural Economics*, 62:584.

Peck, A. E. (1975): "Hedging and Income Stability: Concepts, Implications and an Example," *American Journal of Agricultural Economics,* 57:410–419.

Rutledge, D. J. S. (1972): "Hedgers' Demand for Futures Contracts: A Theoretical Framework with Applications to the United States Soybean Complex," *Stanford Food Research Institute Studies in Agricultural Economics, Trade and Development,* 11:237–256.

Samuelson, P. A. (1965): "Proof That Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review,* 6:41–49.

Stevenson, R. A., and Bear, R. M. (1970): "Commodity Futures: Trends or Random Walks?," *The Journal of Finance*, 25:65–81.

Tomek, W. G. and Gray, R.W. (1970): "Temporal Relationships Among Prices on Commodity Futures Markets: Their Allocative and Stabilizing Roles," *American Journal of Agricultural Economics,* 52:372–380.

Vaccaro, R. J., and Fu Li (1987): "A State-Space Approach for Obtaining Spectral Models from Nonpositive Covariance Models," Proceedings IEEE Conference on Decision and Control, 3:1050–1055.

Ward, R.W., and Fletcher, L. B. (1971): "From Hedging to Pure Speculation: A Micro Model of Optimal Futures and Cash Market Positions," *American Journal of Agricultural Economics,* 53:71–78.